

Classification of Modern Greek Free Subject Verb MWEs

Niki Samaridi¹ and Stella Markantonatou²

¹National and Kapodistrian University of Athens, nsamaridi@gmail.com

²Institute for Language and Speech Processing/ "Athena" RIC, marks@ilsp.gr

PARSEME 3rd general meeting, Frankfurt, Germany, 8-9 September 2014 – WG1

The corpus

A corpus of 2569 Free Subject Verb MWEs:

- Study of their syntactic structures
- Suggestion of a simple and exhaustive classification of them



The development of a parsing system for Modern Greek (MG) free subject Verb MWEs (Samaridi and Markantonatou, 2014)

Our Proposal

MWEs show a degree of semantic non-compositionality and/or a degree of morpho-syntactic inflexibility and vary wrt **the degree of fixedness between the verb and its complements.**

Grammatical framework independent classification of free subject verb MWEs (five categories)

1. verb + one inflexible XP [fixed (i, ii), semi-fixed (iii, iv)]

- i) η Μαρία ήπιε [NP το αμίλητο νερό]
the Mary drank [NP the non-speaking water]
'Mary is silent.'
- ii) η Μαρία έπεσε [PP από τα σύννεφα]
the Mary fell [PP from the clouds]
'Mary was taken by surprise.'
- iii) *η Μαρία_j έφτασε [AP τελευταία_j και καταϊδρωμένη_j]
the Mary_j arrived [AP last_j and battening_j]
'Mary arrived late.'

*The adjectives ('τελευταία', 'καταϊδρωμένη') agree in gender, number and case with the subject.

- iv) *η Μαρία_j κάνει_j [XP πως δεν βλέπει_j]
the Mary_j does_j [XP that not see_j]
'Mary ignores something.'

*The verb 'κάνει' behaves as a control verb and the subject 'η Μαρία' is the controller.

2. verb + one [XP with inflexible head + XPgen]

- i) η Μαρία_j τρώει [NP [τα νύχια] της_j]
the Mary_j eats [NP [the nails] hers_j]
'Mary is nervous.'
- ii) *η Μαρία άναψε [NP [τα λαμπάκια] του Κώστα] → η Μαρία [XPgen του] άναψε [NP τα λαμπάκια]
the Mary lit [NP [the lights] the Costas] → the Mary [XPgen he] lit [NP the lights]
'Mary irritated Costas.'
- iii) η Μαρία πιάστηκε [PP [στα δίχτυα] του Κώστα]
the Mary was caught [PP [in the nets] the Costas]
'Mary was trapped by Costas.'
- iv) *η Μαρία μπήκε [PP [στο ρουθούνι] του Κώστα] → η Μαρία [XPgen του] μπήκε [PP στο ρουθούνι]
the Mary entered [PP [in the nostril] the Costas] → the Mary [XPgen he] entered [PP in the nostril]
'Mary annoyed Costas.'

*In (ii,iv) the possessive genitive 'του Κώστα' may be replaced by the weak personal pronoun 'του' that precedes the verb.

Diagnostics of constituency

When the degree of inflexibility is high, the diagnostics that apply are the following:

- **Fixed word order** (2.c: i): η Μαρία έκανε [[την τρίχα] [τριχιά]] → *η Μαρία έκανε [[τριχιά] [την τρίχα]].
- **No external elements intervene between the components** (2.c: iv): η Μαρία τραβάει [[το σκοινί] [μέχρι να σπάσει]] → *η Μαρία τραβάει [το σκοινί] πάλι [μέχρι να σπάσει].

When there is a high degree of flexibility, the following diagnostics also apply:

- **Reordering** (2.d: ii): η Μαρία έριξε [NP το μπαλάκι] [PP στον Κώστα] → η Μαρία έριξε [PP στον Κώστα] [NP το μπαλάκι].
- **External modifiers of the entire range of MWE** (2.b: i): η Μαρία τρώει [NP [τα νύχια] της_j] → η Μαρία_j τρώει πάλι [NP [τα νύχια] της_j] = the Mary_j eats again [NP [the nails] hers_j].

Types of MWEs

- **Lexicalized and Institutionalized Expressions** (Bauer, 1983), on the basis of their semantic compositionality.
- **Fixed, Semi-Fixed and Syntactically Flexible Expressions** (Sag et al., 2001), on the basis of their flexibility.

The previous classification of free subject verb MG MWEs

by Fotopoulou (1993) describes a wide range of verb MWE syntactic structures, but

- it includes a large number of classes (13 in all)
- it adheres to the Lexicon-Grammar framework
- it does not cover all the syntactic structures and the linguistic phenomena associated with them, such as **subject control phenomena**: e.g. Η Ελένη_j έφαγε_j τον κόσμο να τον βρει_j

the Helen_j ate_j the world to him find_j
'Helen searched everywhere to find him'

3. verb + two inflexible XPs [fixed (i-iv), semi-fixed (v)]:

The order of constituents is usually fixed:

- i) η Μαρία έκανε [XP [NP την τρίχα] [NP τριχιά]]
the Mary made [XP [NP the hair] [NP rope]]
'Mary exaggerated things.'
- ii) η Μαρία έβγαλε [XP [NP το φίδι] [PP από την τρύπα]]
the Mary pulled-out [XP [NP the snake] [PP from the hole]]
'Mary found a solution for a difficult problem.'
- iii) η Μαρία πέφτει [XP [PP από έκπληξη] [PP σε έκπληξη]]
the Mary falls [XP [PP from surprise] [PP to surprise]]
'Mary is taken by surprise all the time.'
- iv) η Μαρία τραβάει [[το σκοινί] [μέχρι να σπάσει]]
the Mary pulls [[the rope] [until to break]]
'Mary crosses the line.'
- v) *η Μαρία_j έκανε_j [NP μαύρα μάτια] [XP να τον δει_j]
the Mary_j made_j [NP black eyes] [XP to him see_j]
'It took Mary a lot of time to meet him.'
- *The verb 'έκανε' behaves as a control verb and the subject 'η Μαρία' is the controller.

4. verb + one inflexible XP + one free XP:

- i) η Μαρία έφαγε [NP τον Κώστα] [NP λάχανο]
the Mary ate [NP the Costas] [NP cabbage]
'Mary beat Costas.'
- ii) η Μαρία έριξε [NP το μπαλάκι] [PP στον Κώστα]
the Mary threw [NP the ball] [PP to the Costas]
'Mary put the blame on Costas.'
- iii) η Μαρία ήρθε [PP στα χέρια] [PP με τον Κώστα]
the Mary came [PP into hands] [PP with the Costas]
'Mary had a (psysical) fight with Costas.'

5. verb + three XPs (at least one inflexible):

- i) η Μαρία έδωσε [PP στον Κώστα] [NP λαβή] [PP για σχόλια]
the Mary gave [PP to the Costas] [NP handle] [PP for comments]
'Mary gave a handle to Costas for comments.'
- ii) η μητέρα έκανε [NP την Μαρία] [[NP τόπι] [PP στο ξύλο]]
the mother made [NP the Mary] [[NP ball] [PP into beating]]
'Mother beat Mary brutally.'

References

- Sag, Ivan A., Baldwin, Timothy, Bond, Francis, Copestake, Ann, and Flickinger, Dan. 2001. Multi-word Expressions: A Pain in the Neck for NLP. Paper presented at *LinGO Working Papers*, Stanford University, CA.
- Fotopoulou, Aggeliki. 1993. *Une Classification des Phrases a Complements Figes en Grec Moderne*. Doctoral Thesis, Universite Paris VIII.
- Samaridi, N. and Markantonatou, S. 2014. Parsing Modern Greek verb MWEs with LFG/XLE grammars. *The 10th Workshop on Multiword Expressions (MWE 2014)*, Workshop at EACL 2014 (Gothenburg, Sweden), April 26-27, 2014